

Maximal Stabilities of Reversible Two-State Proteins[†]Sandeep Kumar,[‡] Chung-Jung Tsai,[§] and Ruth Nussinov*^{*,§,||}

Laboratory of Experimental and Computational Biology, National Cancer Institute at Frederick, Building 469, Room 151, Frederick, Maryland 21702, Intramural Research Support Program-SAIC, Laboratory of Experimental and Computational Biology, National Cancer Institute at Frederick, Building 469, Room 151, Frederick, Maryland 21702, and Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

Received December 13, 2001; Revised Manuscript Received March 6, 2002

ABSTRACT: The hydrophobic effect is the major force driving protein folding. Around room temperature, small organic solutes and hydrophobic amino acids have low solubilities in water and the hydrophobic effect is the strongest. These facts suggest that globular proteins should be maximally stable around room temperature. While this fundamental paradigm has been expected, it has not actually been shown to hold. Toward this goal, we have collected and analyzed experimental thermodynamic data for 31 proteins that show reversible two-state folding \rightleftharpoons unfolding transitions at or near neutral pH. Twenty-six of these are unique, and 20 of the 26 are maximally stable around room temperature irrespective of their structural properties, the melting temperature, or the living temperatures of their source organisms. Their average temperature of maximal stability is 293 ± 8 K (20 ± 8 °C). These proteins differ in size, fold, and number of domains, hydrophobic folding units, and oligomeric states. They derive from the cold-loving psychrophiles, from mesophiles, and from thermophiles. Analysis of the single-domain proteins present in this set shows that the variations in their thermodynamic parameters are correlated in a way which may explain the adaptation of the proteins to the living temperatures of the organisms from which they derive. The average energetic contribution of the individual amino acids toward protein stability decreases with an increase in protein size, suggesting that there may be an upper limit for protein maximal thermodynamic stability. For the remaining proteins, deviation of the maximal stability temperatures from room temperature may be due to greater uncertainties in their heat capacity change (ΔC_p) values, a weaker hydrophobic effect, and/or a stronger electrostatic contribution.

Protein stability varies with temperature, the presence of chemical denaturants, changes in the solvent, pH, salt concentration, and buffer composition. For two-state proteins, the transition from the native (N) to the denatured (D) state is reversible, with no intermediate states. Their folding \rightleftharpoons unfolding profiles show a high degree of cooperativity. Such proteins are often small with single domains, are usually stable over a temperature range, and have a constant (>0) heat capacity difference between the native and denatured states within this range. To study the temperature-dependent thermodynamic stability of a two-state protein, its stability

curve can be plotted using the Gibbs–Helmholtz equation (1, 2)

$$\Delta G(T) = \Delta H_G(1 - T/T_G) - \Delta C_p[T_G - T + T \ln(T/T_G)] \quad (1)$$

where $\Delta G(T)$ ¹ is the Gibbs free energy change between the denatured (D) and native (N) states of the protein at a given temperature T , ΔH_G is the enthalpy change between the two states at the melting temperature (T_G), and ΔC_p is the heat capacity change between the two states. A positive $\Delta G(T)$ indicates that the native state is more stable than the denatured state of the protein at a temperature T . Equation 1 provides the value of $\Delta G(T)$ at T if we know the experimentally determined T_G , ΔH_G , and ΔC_p values. These

[†] The research of R.N. in Israel has been supported in part by the Magnet grant and by the Center of Excellence in Geometric Computing and its Applications, funded by the Israel Science Foundation, by a Ministry of Science grant, and by Tel Aviv University Basic Research grants. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract NO1-CO-12400.

* To whom correspondence should be addressed: NCI-FCRF Building 469, Room 151, Frederick, MD 21702. Telephone: (301) 846-5579. Fax: (301) 846-5598. E-mail: ruthn@ncifcrf.gov.

[‡] Laboratory of Experimental and Computational Biology, National Cancer Institute at Frederick.

[§] Intramural Research Support Program-SAIC, Laboratory of Experimental and Computational Biology, National Cancer Institute at Frederick.

^{||} Tel Aviv University.

¹ Abbreviations: T_G , heat denaturation (melting) temperature; T'_G , cold denaturation temperature; T_S , temperature of maximal protein stability; ΔH_G , molar enthalpy change between the native and denatured states of a protein at T_G ; ΔS_G , molar entropy change between the native and denatured states of a protein at T_G ; ΔC_p , molar heat capacity change between the native and denatured states of a protein; $\Delta G(T)$, molar Gibbs free energy change between the native and denatured states of a protein at a temperature T ; $\Delta G(T_S)$, molar Gibbs free energy change between the native and denatured states of a protein at T_S ; DSC, differential scanning calorimetry; CD, circular dichroism; r , linear correlation coefficient; HFU, hydrophobic folding unit.

are obtained from thermal and chemical denaturation using spectroscopic (CD and FI) and calorimetric (DSC) techniques.

Examination of a typical protein stability curve shows that each two-state protein shows a characteristic temperature at which its thermodynamic stability is maximum (1). This temperature is given by

$$T_S = T_G \exp[-\Delta H_G / (T_G \Delta C_p)] \quad (2)$$

Every two-state protein has two transition temperatures where $\Delta G(T) = 0$ (1, 2). These are the heat (T_G) and cold (T'_G) denaturation temperatures. Heat denaturation is driven by an increase in entropy (3) and cold denaturation by a decrease in enthalpy (2, 4, 5). Consideration of a two-state model of the water structure (6, 7) aids in understanding the microscopics of protein denaturation (8, 9). In such a two-state water model, the first state is the enthalpically favored, icelike tetrahedrally connected hexagonal hydrogen bonding form with optimal hydrogen bond networks. The second state is the entropically favorable, highly fluctuating liquid form. The proportions of these hydrogen bonding types show a temperature-dependent gradient. At low temperatures, the hexagonal icelike hydrogen bonding type prevails. At high temperatures, the denser liquid type dominates. At any given temperature, the water structure is dynamic with hydrogen bonds continuously formed and broken on a very short time scale (10). In the temperature ranges where $T \leq T'_G$ and $T \geq T_G$, where the denatured states of the protein are energetically favorable, the significant changes in the water structure drive protein denaturation (8, 9, 11). Thus, the hydrophobic effect is a function of temperature, via changes in the water structure.

Even though the relative contributions made by the hydrophobic effect and the electrostatic interactions toward protein stability are still controversial (12, 13), the hydrophobic effect is considered to be the major driving force in protein folding and binding (14, 15). The hydrophobic effect similarly explains clathrate formation of small organic solutes in water. The hydrophobic effect is the strongest around 20 °C. The contribution of the hydrophobic effect to protein stability decreases at lower and higher temperatures (16). At room temperature, the enthalpies of dissolution of nonpolar molecules in water are negative with their magnitudes proportional to the accessible surface area of the solute molecules (17). The entropy of a nonpolar substance in water is also negative at room temperature, and its magnitude decreases with an increase in temperature (17). Additionally, small organic solutes have minimum solubilities around room temperature. For example, the solubility of hexane is only 2.0×10^{-6} mole fractions, and for benzene, it is 4.01×10^{-4} mole fractions at 25 °C (17). Figure 1 shows plots of the solubilities of seven hydrophobic amino acids in water at different temperatures. All have low solubilities in the temperature range of 0–30 °C. The solubilities of several naturally occurring amino acids (Ala, Leu, Asp, Trp, Tyr, Cys, Glu, Ile, and Val) remain low and almost constant in this 0–30 °C range (18). These observations imply that a reversible two-state globular protein with a sufficiently large hydrophobic core should be maximally stable around room temperature.

Rees and Robertson (19) have predicted the temperature of maximal stability for proteins to be ~ 283 K (10 °C) on

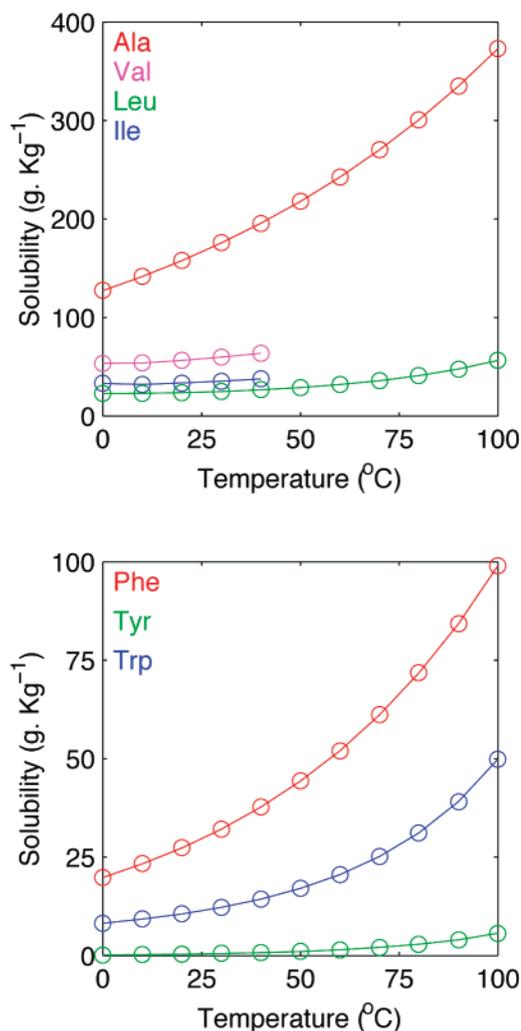


FIGURE 1: Plots showing solubilities (grams per kilogram) of seven hydrophobic amino acids in proteins in water as a function of temperature (degrees Celsius). Around room temperature, the solubilities of amino acids are low and remain more or less constant in the range of 0–30 °C for Val, Leu, Ile, Trp, and Tyr. For Ile and Val, the solubility data are available up to only 40 °C. The data for these plots were taken for the L-forms of the amino acids from ref 18.

the basis of a quadratic function approximation to the Gibbs–Helmholtz equation. They used parametrization derived from a survey of protein thermodynamic data by Robertson and Murphy (20). Using the same data set, they found the proteins to be maximally stable at 285 ± 19 K. Ganesh et al. (21) have observed that the temperatures of maximal stabilities for mesophilic proteins range between -25 and 35 °C.

We have collected experimental thermodynamic data on 31 proteins that show highly reversible two-state folding \rightleftharpoons unfolding transitions at or near neutral pH. Twenty-six of these proteins are unique, and 20 of the 26 are maximally stable around room temperature with the average temperature of maximal stability being 293 ± 8 K (20 ± 8 °C, range of 1–34 °C). This occurrence of maximal stability around room temperature appears to be independent of the structural details (number of residues, folds, domains, hydrophobic folding units, and oligomeric states) of these proteins, their melting temperatures, and the living temperatures of their source organisms (psychrophiles, mesophiles, and thermophiles). We have also obtained correlations among different thermo-

dynamic parameters across 12 single-domain proteins. These correlations suggest that higher melting temperatures could be achieved by upshifting and broadening the protein stability curves and that there may be an upper limit to maximal thermodynamic protein stability.

MATERIALS AND METHODS

Data Collection. We have performed a Pubmed search for reports containing experimental thermodynamics data on proteins. The search was supplemented by thermodynamic data contained in the ProTherm database (22) and that compiled by Pfeil (23). Our aim has been to collect thermodynamic data on proteins that exhibit a reversible two-state folding \rightleftharpoons unfolding transition at or near neutral pH (pH range of 6–8). By restricting the pH values for the folding \rightleftharpoons unfolding transition to neutral pH, we are able to separate the effect of temperature on protein stability from that of the hydrogen ion concentration in solution. Further, the thermodynamic data collected here are in the absence of denaturing agents such as urea or guanidinium hydrochloride. Circular dichroism (CD) and differential scanning calorimetry (DSC) are the techniques most widely used in studies of protein denaturation. Analysis of protein denaturation and renaturation profiles obtained from these studies has been standardized. These techniques often yield relatively accurate values of ΔH_G , T_G , and ΔC_p . Hence, we have restricted ourselves to data reported by experiments using either CD, DSC, or both. Data from other techniques such as fluorescence and NMR were taken only when accompanied by data from at least one of the two techniques (CD or DSC). Table 1 contains data on thermodynamic parameters (T_G , ΔH_G , and ΔC_p), pH, buffer, and salt concentration, on the 31 proteins in our database. The following criteria were used for choosing proteins in our data set.

(i) **Evidence for Two-State Folding Behavior.** We have accepted the claim of the original authors on the two-state nature of the folding behavior of the proteins in our database. This claim is supported either by the presence of isodichroic point(s) in the CD spectra recorded at the different temperatures in the transition region, by a single peak in the DSC scan, or by both. The extent to which a protein follows the two-state mechanism can be measured by the cooperativity ratio, R :

$$R = \Delta H^{\text{cal}} / \Delta H^{\text{van'tHoff}} \quad (3)$$

where ΔH^{cal} is the enthalpy change for unfolding determined by DSC. This enthalpy change value is model-independent. $\Delta H^{\text{van'tHoff}}$ is determined from thermal denaturation experiments using CD spectroscopy. Alternatively, it is calculated by (24)

$$\Delta H^{\text{van'tHoff}} = (4RT_G^2 C_{p,\text{max}}) / \Delta H^{\text{cal}} \quad (4)$$

where $C_{p,\text{max}}$ is the maximum excess heat capacity at T_G and R is the universal gas constant. The calculation of the $\Delta H^{\text{van'tHoff}}$ value assumes a two-state folding model. Thus, a value of R close to unity indicates the validity of a two-state folding model for a monomeric protein (24, 25). In our database, we have selected proteins where $0.9 \leq R \leq 1.10$, wherever these statistics are available (24 of 31 proteins).

For 17 proteins in our database, the values of R lie within the range of 0.95–1.05. We have also checked for concentration, calorimeter scan rate independence of the protein folding \rightleftharpoons unfolding transition, and whether ΔC_p is constant. For 28 of the 31 proteins, the pH values are within the range of 6.5–7.5, and for 17 proteins, the transitions were recorded at pH 7.0.

(ii) **Reversibility of the Protein Folding \rightleftharpoons Unfolding Transition.** The reversibility of the protein folding \rightleftharpoons unfolding transition can be measured by the reproducibility of the DSC or CD scans on the same protein sample. All proteins in our database exhibit $\geq 90\%$ reversibility. For 20 proteins, the reversibility is $\geq 95\%$. The accuracy in the determination of the three thermodynamic parameters used in this study is indicated by the standard deviations about the mean values, wherever available.

The data in the literature are frequently reported in SI and non-SI units. For uniformity, we use calorie as the unit of energy. The conversion factor is 1 cal = 4.184 J.

Protein Stability Curves. Using the Gibbs–Helmholtz equation (eq 1), we have computed and plotted the protein stability curves for the 31 proteins (Figure 2). Additional derived thermodynamic characteristics include the entropy change at the melting temperature ($\Delta S_G = \Delta H_G / T_G$), the temperature of maximal protein stability (T_S), the maximal protein stability [$\Delta G(T_S)$], and the protein stability at room temperature. Analyses were performed on 26 unique proteins: α -amylase from psychrophile *Alteromonas haloplacis*, DNA-binding protein Sso7d from hyperthermophile *Sulfolobus solfataricus*, *Thermotoga maritima* cold shock protein (TmCsp), *Bacillus subtilis* histidine phosphocarrier protein (BsHPr), λ repressor_{6–85}, barstar, ribonuclease Sa, ribonuclease T₁, ribonuclease A, binase, maltose binding protein (MBP), odorant binding protein (OBP), *Escherichia coli* repressor of primer (ROP), staphylococcal nuclease, ferricytochrome *b*₅₆₂, Arc repressor, *T. maritima* glutamate dehydrogenase domain II (TmGDH domain II), histone H2A–H2B dimer, C_L fragment (residues 109–212) of Ig λ , rat thyroid transcription factor 1 homeodomain (TTF-1HD), Kunitz type soybean trypsin inhibitor (STI), apoflavodoxin, thioredoxin, Met J dimer, agglutinin (ASAI), and the activation domain of human procarboxypeptidase A2 (ADA2h).

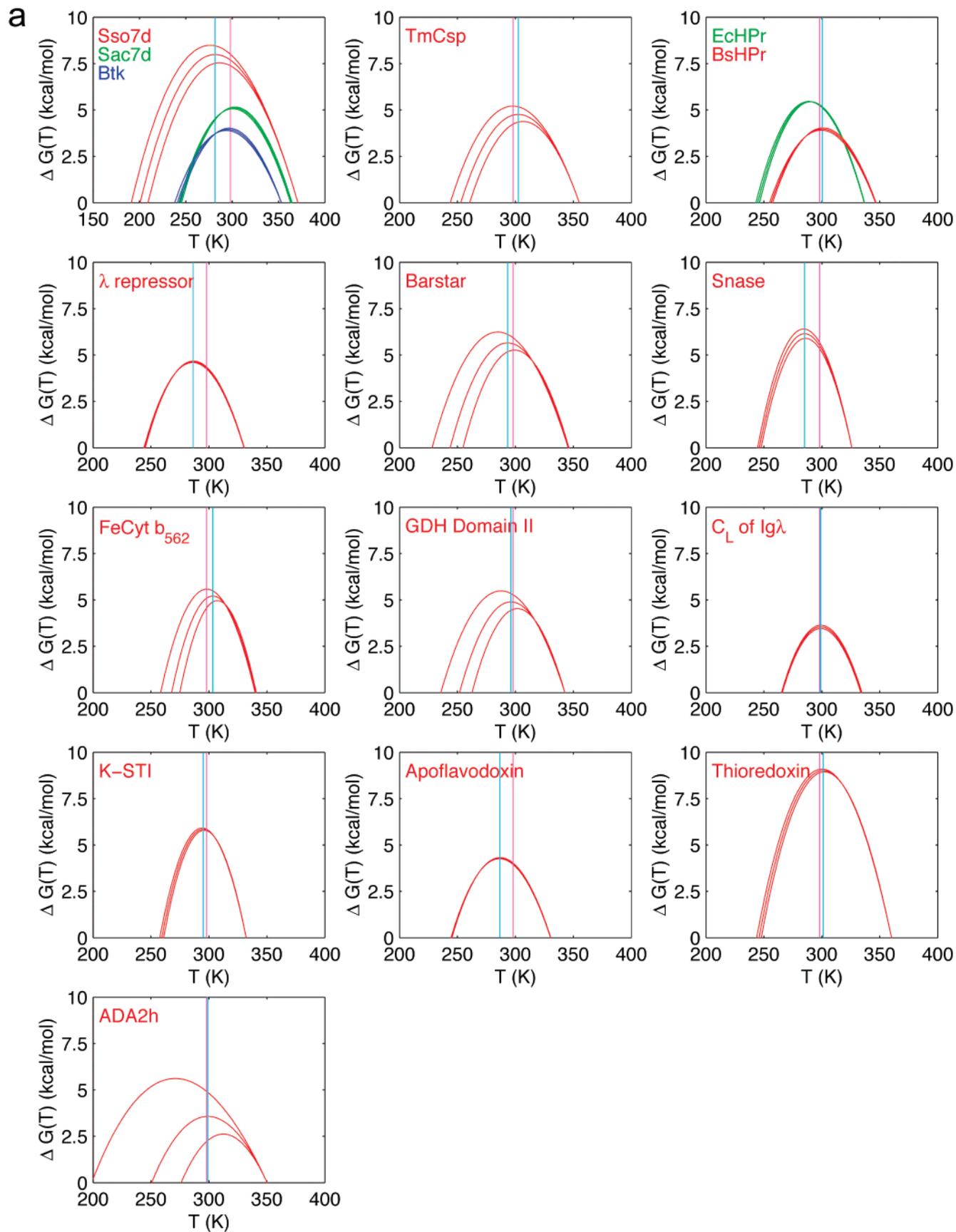
Three-Dimensional Structures, Hydrophobicity, Hydrophobic Folding Units, and Domains. We have taken three-dimensional (3D) structures determined by either crystallography or NMR for proteins from the Protein Data Bank (26). These structures are available for 27 (of 31) proteins in our database. The hydrophobicity of a protein was calculated as the fraction of the buried nonpolar area out of the total nonpolar area, computed by using the methods described previously (27, 28). A hydrophobic folding unit (HFU) is a part of the protein structure with significant buried hydrophobic core, capable of an independent thermodynamically stable existence (27, 28). The HFUs in each protein were identified using the procedure described by Tsai and Nussinov (27, 28). Table 2 lists these structures, the number of HFUs, and the values of hydrophobicity for each protein. We have used the PDBsum (<http://www.biochem.ucl.ac.uk/bsm/pdbsum/index.html>) database for retrieving the CATH domain assignments for our proteins.

Table 1: Proteins That Exhibit a Reversible Two-State N \rightleftharpoons D Transition at or near Neutral pH^a

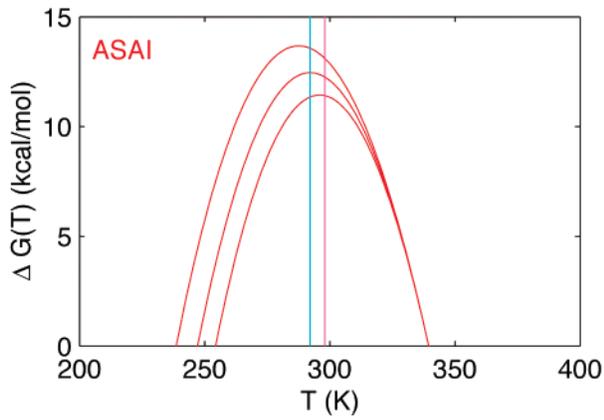
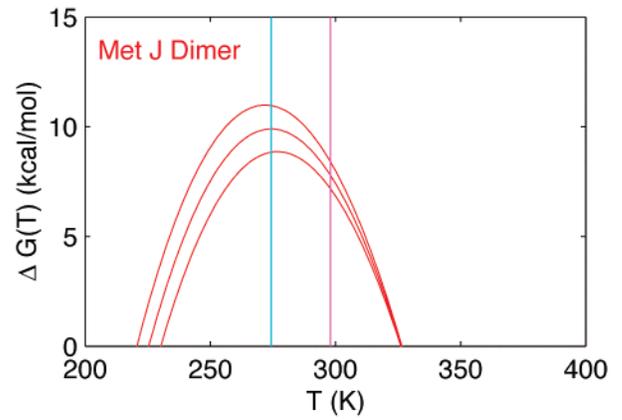
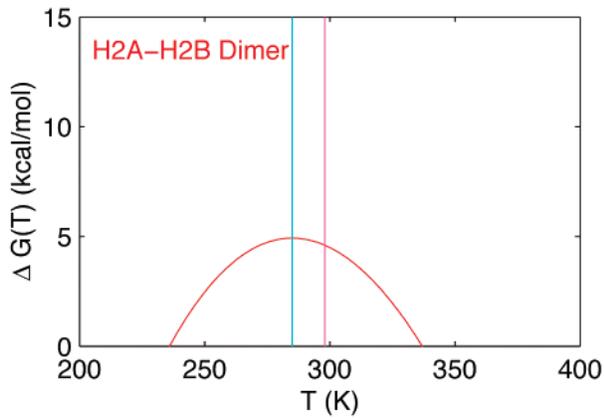
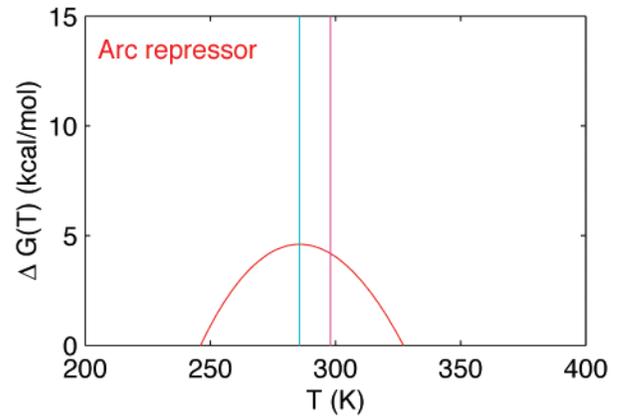
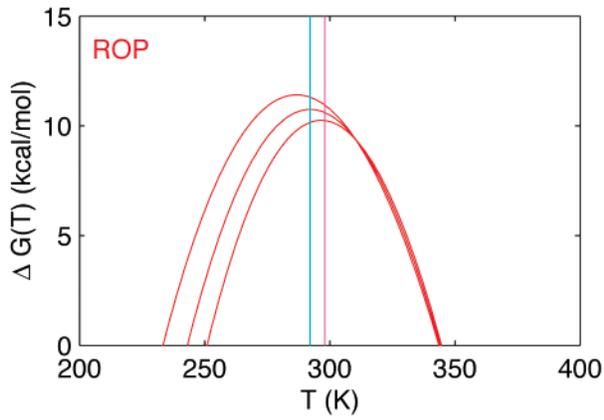
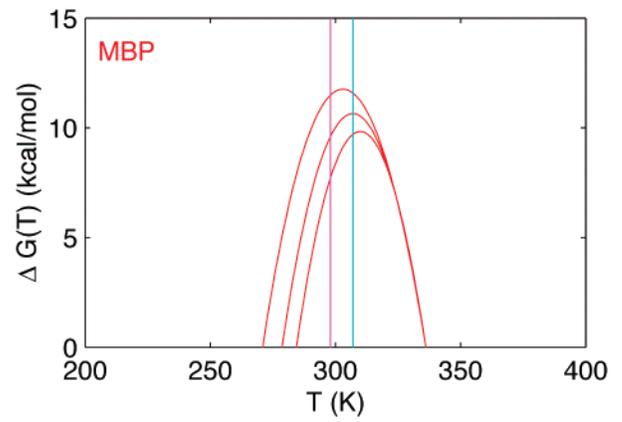
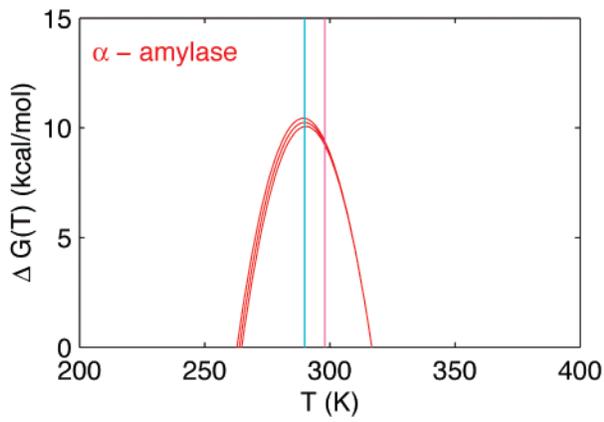
protein	source	N_{res}	T_G (°C)	ΔH_G (kcal/mol)	ΔC_p (kcal mol ⁻¹ K ⁻¹)	pH	buffer and salt concentration	exp. tech.	rev. (%)	R
α -amylase ^b (two domains)	<i>A. haloplanctis</i> (psychrophile)	453	43.7	238	8.47 \pm 0.16	7.2	30 mM Mops, 5 mM EGTA	DSC, CD, FI	~99	1.06
Sso7d ^c (one domain)	<i>S. solfataricus</i> (thermophile)	62	97.8	63.4	0.62 \pm 0.04	6.5	Tris/Cl ⁻ /Mes/K ⁺	DSC, CD	90–98	0.92–0.95
Sac7d ^d (one domain)	<i>Sulfolobus acidocaldarius</i> (thermophile)	66	90.7 \pm 0.8	58.3 \pm 1.0	0.86 \pm 0.02	7.0	0.3 M KCl	DSC, CD	90	1.03 \pm 0.12
Btk ^e (one domain)	<i>Homo sapiens</i>	67	80	46.8 \pm 1.9	0.74 \pm 0.05	6.0		DSC, CD	>93	0.91 \pm 0.05
TmCsp ^f (one domain)	<i>T. maritima</i> (thermophile)	66	82	62.6	1.1 \pm 0.1	7.0	cacodylate	DSC	97 \pm 2	0.94 \pm 0.06
EcHPr ^g (one domain)	<i>E. coli</i>	85	63.6 \pm 0.1	75.8 \pm 1.2	1.49 \pm 0.05	7.0	10 mM KP _i	CD	\geq 95	
BsHPr ^h (one domain)	<i>B. subtilis</i>	87	73.4 \pm 0.2	58.1 \pm 1.7	1.17 \pm 0.05	7.0	10 mM KP _i	CD	\geq 95	
λ repressor 6-85 ⁱ (one domain)	λ phage	80	57.2 \pm 0.1	68.0 \pm 1.0	1.44 \pm 0.03	8.0	100 mM NaCl, 20 mM KD ₂ PO ₄ , 99% D ₂ O	CD, NMR	>90	
barstar ^j (one domain)	<i>Bacillus amyloliquefaciens</i>	90	72.7 \pm 0.3	72.4 \pm 3.6	1.27 \pm 0.24	7.4	50 mM Na ₃ PO ₄ , 1 mM EDTA, 10 mM DTT	DSC	>90	1.05 \pm 0.1
Rnase Sa ^k (one domain)	<i>Streptomyces aurefaciens</i>	96	48.4 \pm 0.3	97.4 \pm 4.9	1.52 \pm 0.09	7.0	30 mM MOPS	DSC, CD	>95	0.99 \pm 0.08
Rnase Sa2 ^l (one domain)	<i>St. aurefaciens</i>	97	41.1 \pm 0.3	68.4 \pm 3.4	1.27 \pm 0.06	7.0	30 mM MOPS	DSC, CD	>95	1.00 \pm 0.03
Rnase Sa3 ^m (one domain)	<i>St. aurefaciens</i>	99	47.2 \pm 0.3	93.6 \pm 4.7	1.57 \pm 0.11	7.0	30 mM MOPS	DSC, CD	>95	0.96 \pm 0.04
Rnase T ₁ ⁿ (one domain)	<i>Aspergillus oryzae</i>	104	48.9 \pm 0.1	95.5 \pm 0.9	1.59	7.0	30 mM PIPES	DSC	~100	~1.00
Rnase A ^o (one domain)	bovine pancreas	110	61.8	102.3	1.15 \pm 0.09	7.0	10 mM MOPS	DSC, ITC	~100	1.02 \pm 0.02
binase ^p (one domain)	<i>Bacillus intermedius</i> 7P	109	57.6	118.5	0.85 \pm 0.23	7.0	10 mM glycine buffer	DSC	~100	1.05
maltose binding protein ^q (two domains)	<i>E. coli</i>	370	63.0	241.4 \pm 7.2	7.9 \pm 1.2	7.4	CGH10	DSC, CD	91	~1.0
odorant binding protein ^r (dimer with one domain per monomer)	porcine	149 per monomer	69.2 \pm 0.3	93.5 \pm 4.1	0.95 \pm 0.33	6.6	10 mM phosphate, 1 mM EDTA	DSC, CD	~100	0.997 \pm 0.01
ROP ^s (dimer with one domain per monomer)	<i>E. coli</i>	63 per monomer	71.0 \pm 0.5	138.6 \pm 4.8	2.46 \pm 0.31	6.0	10 mM Na ₃ PO ₄ , 10 mM Na ₂ SO ₄ , 1 mM EDTA	DSC, CD	~100	1.07
Snase ^t (one domain)	<i>Staphylococcus aureus</i>	149	52.8	96 \pm 2	2.2	7.0	0.1 M NaCl, 25 mM Na ₃ PO ₄	DSC, FI	99	1.05
cytochrome <i>b</i> ₅₆₂ (ferri) ^u (one domain)	<i>E. coli</i>	106	67.2 \pm 0.5	94 \pm 5	2.4 \pm 0.4	7.0	100 mM Na ₃ PO ₄ , 0.1 mM EDTA	CD	>90	
Arc repressor ^v (dimer with one domain per monomer)	phage P22	53 per monomer	54.0	71.0	1.6	7.3	10 mM Tris, 100 mM KCl	CD	90	
glutamate dehydrogenase domain II ^w (one domain)	<i>T. maritima</i> (thermophile)	150	69.6	70.2 \pm 4.0	1.4 \pm 0.3	7.5	20 mM Na ₃ PO ₄ , 100 μ M DTT, 100 μ M EDTA	DSC, CD	98	0.91
histone H2A–H2B dimer ^x (one domain per monomer)	chicken erythrocyte	129 (H2A), 126 (H2B)	63.8	62.1	1.1	7.5	0–140 mM NaCl, 1 mM EDTA, 10 mM imidazole	DSC, CD	99	0.92
C _L fragment (residues 109–212) of Ig λ ^y (one domain)	<i>H. sapiens</i>	104	61 \pm 0.5	66.5 \pm 0.9	1.8	7.5	0.01 M Na ₃ PO ₄ , 0.15 M NaCl	CD, FI	~100	
thyroid transcription factor 1 homeodomain ^z (TTF-1HD) (one domain)	rat	61	42.8	25.41	0.34	7.5	10 mM NH ₃ HPO ₃	CD, FI	~100	
Kunitz type soybean trypsin inhibitor (STI) ^{aa} (one domain)	soybean	181	59.0	102.5 \pm 1.4	2.6 \pm 0.1	7.0	20 mM phosphate buffer	DSC	~100	0.91 \pm 0.04
flavodoxin (apo form) ^{bb} (one domain)	<i>Anabaena</i> PCC 7119	168	57.3 \pm 0.1	63.1 \pm 0.7	1.34 \pm 0.02	7.0	50 mM Na ₃ PO ₄ buffer	DSC, CD, NMR	>90	1.00 \pm 0.04

thioredoxin ^{cc} (one domain)	<i>E. coli</i>	108	87.0	106.9 ± 1.1	1.66 ± 0.05	7.0	0.01 M phosphate buffer, 0.1 M NaCl	DSC	>90	~1.0
MetJ ^{dd} (dimer with one domain per monomer)	<i>E. coli</i>	104 per monomer	53.2 ± 0.2	120.7 ± 6.7	2.13	7.0	25 mM KPi, 100 mM KCl, 1 mM DTT	DSC	~100	0.97 ± 0.09
agglutinin (ASAI) ^{ee} (dimer with one domain per monomer)	<i>Allium sativum</i> (garlic)	106, 109	66.4	174.1	3.41 ± 0.32	7.5	20 mM PBS	DSC, CD	≥90	1.01
ADA2h ^{ff} (one domain)	human	80	77.0	47.6	0.86 ± 0.33	7.0	50 mM phosphate buffer	DSC, CD	>90	1.0 ± 0.05

^a A summary of the thermodynamic data for proteins in our database. For each protein, its name, number of domains, oligomeric state, source organism, number of residues (N_{res}), melting temperature (T_G), enthalpy change at the melting temperature (ΔH_G), heat capacity change (ΔC_p), pH, buffer composition and salt concentration, experimental technique(s) used to monitor the N \rightleftharpoons D transition, percent reversibility (rev.) of the transition, and cooperativity ratio ($R = \Delta H^{cal}/\Delta H^{van'tHof}$) of the transition (wherever available) are presented. ^b Psychrophilic α -amylase. Data from ref 49. ^c SH3 domain-containing thermophilic protein Sso7d. Data from refs 40, 50, and 51. ^d SH3 domain-containing thermophilic protein Sac7d. Data from refs 40, 50, and 51. Sac7d exhibits a reversible two-state transition in the pH range of 0–10. ^e SH3 domain-containing mesophilic protein Btk (Bruton's tyrosine kinase). Data from refs 40, 50, and 51. ^f *T. maritima* cold shock protein exhibits a two-state N \rightleftharpoons D transition over the pH range of 3.5–8.5. Data from ref 52. ^g Histidine phosphocarrier protein (Hpr) from a mesophilic organism, *E. coli*. Data from refs 23, 53, and 54. ^h Histidine phosphocarrier protein (Hpr) from a mesophilic organism, *B. subtilis*. Data from refs 23, 53, and 54. ⁱ DNA-binding protein λ repressor. Data from ref 55. ^j Wild-type Barstar exhibits a reversible two-state transition in the pH range of 6.4–8.3. Data from ref 56. ^k The protein contains a single disulfide bond linking its N- and C-termini. Data from ref 31. ^l The protein contains a single disulfide bond linking its N- and C-termini. Data from ref 31. ^m The protein contains a single disulfide bond linking its N- and C-termini. Data from ref 31. ⁿ Ribonuclease T₁ contains two disulfide bonds. The data are for Gln25 Rnase T₁ from ref 57. ^o Ribonuclease A. Data from ref 32. ^p *B. intermedius* ribonuclease. The protein exhibits a reversible two-state transition in the pH range of 4.0–7.0. Data from refs 23 and 58. ^q *E. coli* maltose binding protein (MBP). Data from ref 59. ^r Porcine odorant binding protein is a homodimer. Data from ref 60. ^s Rop is a four-helix bundle. Each subunit contains two α -helices. Data from ref 61. ^t Staphylococcal nuclease. Data from ref 62. ^u Ferricytochrome *b*₅₆₂ is a four-helix bundle. Data from ref 63. ^v Arc repressor is a homodimer. Data from ref 64. ^w The biologically active state of glutamate dehydrogenase from *T. maritima* is a homohexamer. Each subunit contains two domains. The data presented here correspond to the second domain of this enzyme. The thermal unfolding of this domain is highly reversible in the pH range of 5.9–8.0. Data from ref 65. ^x The H2A–H2B dimer is part of the histone octamer (nucleosome core particle). The data used here are for a 14 kDa monomeric unit of the histone H2A–H2B dimer from ref 66. The value of R was calculated from $\Delta H(T^o)$ values at pH 7.5 given in Table 2 of ref 66. ^y A domain consisting of residues 109–212 from the C_L fragment of Ig λ . Data from ref 67. ^z Rat thyroid transcription factor 1 is a homeodomain. Data from ref 33. The value of ΔC_p reported by the authors is 0.08 kcal mol⁻¹ K⁻¹. However, this value appears to be incorrect. The recalculation of ΔC_p from the other data in the paper yields a value of 0.34 kcal mol⁻¹ K⁻¹. We have used this value in our analysis. ^{aa} Kunitz type soybean trypsin inhibitor. Data from ref 68. ^{bb} Flavodoxins are α/β proteins. Data from ref 69. The protein is stable in the pH range of 6–11. For the pH range of 6–9, rev. = 50–60%. But if heating is stopped before the melting temperature, it is >90%. ^{cc} *E. coli* thioredoxin. Data from ref 70. The value of R was judged to be ~1.0 from the DSC scans of thioredoxin shown in Figure 2 of the original reference. ^{dd} MetJ is homodimer. Data from ref 71. ^{ee} ASAI is heterodimer. Data from ref 72. ^{ff} ADA2h is the activation domain of human procarboxypeptidase A2. Data from ref 73.



b



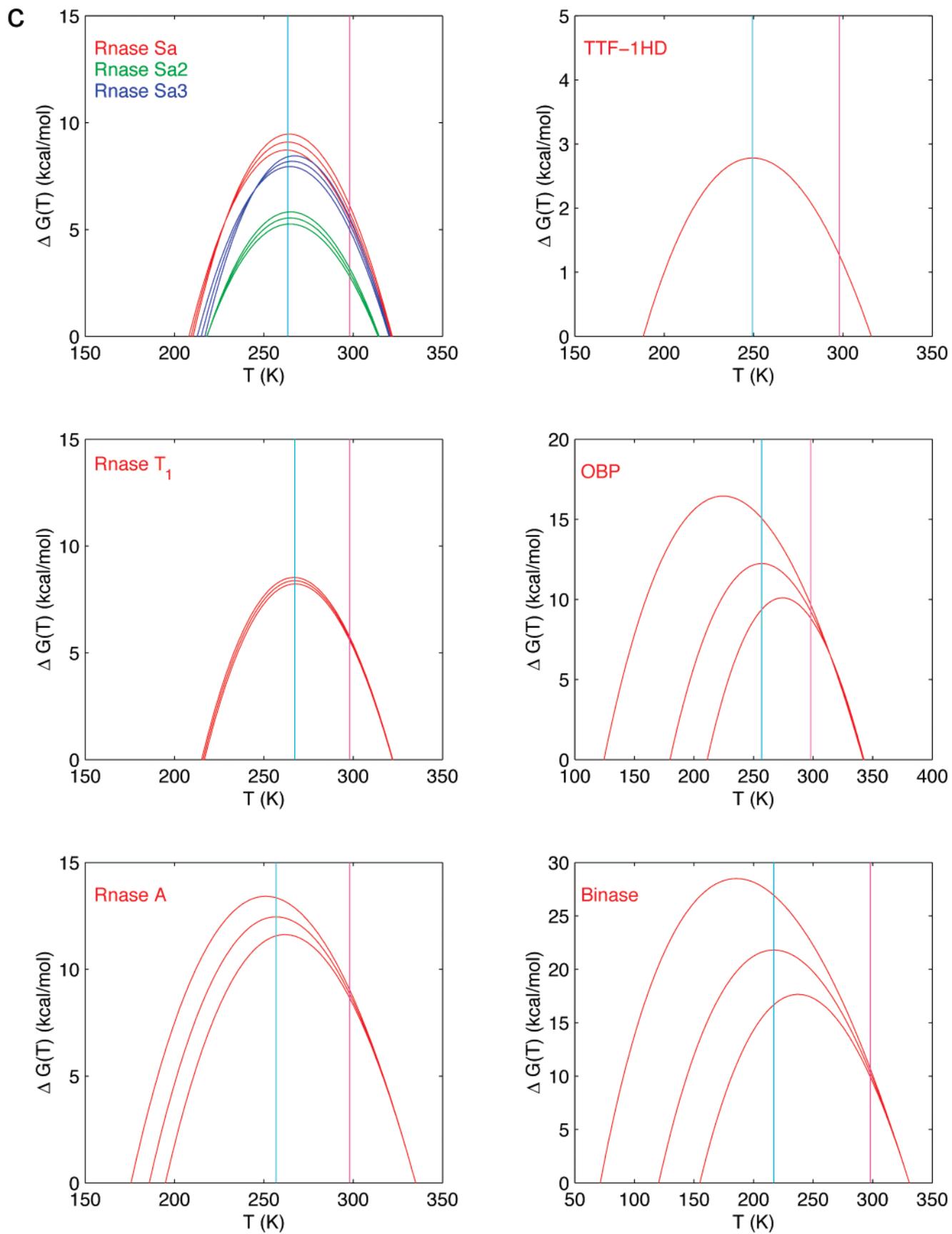


FIGURE 2: Protein stability curves for 31 proteins in our database that show a reversible two state $N \rightleftharpoons D$ transition at or near neutral pH. Twenty-six of these proteins are unique. In each plot, the X -axis represents temperature, T (kelvin), and the Y -axis represents $\Delta G(T)$ (kilocalories per mole). For each stability curve, its lower and upper bounds are also plotted using the available experimental errors in ΔH_G , ΔC_p , and T_G . Additionally, there are two (cyan and magenta) vertical lines. The magenta line is drawn at room temperature (298 K). The cyan line is drawn at the temperature of maximal stability (T_S) of each unique protein. We have presented the protein stability curves of the homologues of a unique protein in the same plot. In such cases, the cyan line corresponds to the T_S of the unique protein selected in our database. (a) Thirteen unique monomeric proteins that contain a single domain. To aid in a direct comparison across the plots, the axes of all plots, except for the Sso7d family, have similar ranges. For the Sso7d family, the X -axis has a greater range. (b) Two monomeric proteins that contain two domains each and five dimeric proteins that contain a single domain per subunit. The axes of all the plots have similar ranges. The protein stability curves in sections a and b present 20 (of 26) unique proteins which are maximally stable around room temperature. (c) Six unique proteins that are maximally stable away from room temperature ($T_S < 273$ K). The ranges of the axes are similar for plots of the Rnase Sa family, Rnase T₁, and Rnase A. However, the ranges of the axes differ for plots of TTF-1HD, binase, and odorant binding protein (OBP).

Table 2: Hydrophobicity Values and Hydrophobic Folding Units (HFUs) in Proteins in Our Database^a

protein	PDB entry	resolution ^b (Å)	hydrophobicity (%)	no. of HFUs
α -amylase	1AQH	2.0	87	3
Sso7d	1SSO		71	1
Sac7d	1SAP		72	1
EcHPr	1POH	2.0	76	1
BsHPr	2HID		80	1
λ -repressor ₆₋₈₅	1LMB	1.8	76	1
barstar	1A19	2.76	80	1
Rnase Sa	1RGG	1.2	75	1
Rnase T ₁	9RNT	1.5	82	1
Rnase A	1AQP	2.0	79	1
binase	1BUJ		77	1
MBP	1OMP	1.8	85	2
OBP	1A3Y	2.25	82	2
ROP	1ROP	1.7	80	2
Snase	1EY0	1.6	79	1
FeCyt b_{562}	1QPU		79	1
Arc repressor	1ARR		76	1
GDH domain II	1B26	3.0	82	1
H2A–H2B dimer	1EQZ	2.5	79	2
C _L fragment	1A8J	2.7	82	1
TTF-1HD	1FTT		66	1
K-STI	1AVU	2.3	80	1
apoflavodoxin	1FTG	2.0	86	1
thioredoxin	2TRX	1.68	79	1
MetJ dimer	1CMB	1.8	81	2
ASAI	1BWU	2.8	84	2
ADA2h	1AYE	1.8	72	1

^a Hydrophobicity and number of hydrophobic folding units (HFU) for proteins in our database for which three-dimensional structures are available in the Protein Data Bank (26). Wherever our database has a domain (e.g., GDH domain II) or fragment [e.g., C_L fragment (residues 109–212) of Ig λ] rather than the full protein, the relevant atomic coordinates were extracted from the PDB. Note that TTF-1HD has a lower hydrophobicity than most of the proteins in our database. ^b The resolution of the structure if determined by X-ray crystallography.

Computation of Linear Correlation Coefficients and Their Statistical Significance. For each pair of parameters (x_1 and y_1 , x_2 and y_2 , ... x_n and y_n) on proteins in our database, we have fitted a least-squares line ($y = mx + c$). The linear correlation coefficient is calculated by (29)

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (5)$$

Our data set can be regarded as a sample from protein populations. Hence, the sampling theory can be used to determine if the correlations observed in our data set are also relevant to proteins in general. We formulate the null hypothesis that the population correlation coefficient (ρ) for a given parameter pair is zero (H_0 , $\rho = 0$), while the linear

correlation coefficient for the same parameter pair is r ($\neq 0$) in our data set. The t value is computed to test the null hypothesis with (29)

$$t = r\sqrt{n-2}/\sqrt{1-r^2} \quad (6)$$

where n is the number of proteins in our data set. The null hypothesis is rejected at 95% ($p < 5 \times 10^{-2}$) or 99.5% ($p < 5 \times 10^{-3}$) levels of confidence if the computed t value is greater than $t_{0.95}$ or $t_{0.995}$ for n proteins (29). Rejection of the null hypothesis for a parameter pair indicates that the two parameters are likely to be correlated with each other in proteins.

For 12 single-domain proteins, the null hypothesis is rejected at the 95% level of confidence if the t value for a parameter pair is > 1.78 ($r \geq |0.5|$). The null hypothesis is rejected at the 99.5% level of confidence if the t value is > 3.06 ($r \geq |0.7|$).

RESULTS

General Observations. Our database consists of 31 proteins or protein domains. Twenty-six proteins are derived from mesophilic organisms; one is from a psychrophile, and four are from thermophiles (Table 1). Twenty-six proteins are unique; i.e., they do not have a homologous protein in our database. Eighteen unique proteins contain single domains; two are monomers with two domains, and six are dimers with a single domain per subunit. While most proteins in our database are small, there are a few exceptions. For example, the α -amylase and MBP monomers, which contain two domains, consist of 453 and 370 amino acids, respectively. Similarly, the dimeric proteins OBP and histone H2A–H2B dimers contain 298 and 255 residues, respectively. Among the single-domain monomeric proteins, STI and apoflavodoxin are also relatively large, containing 181 and 168 residues, respectively.

Figure 2 shows the protein stability curves. For most proteins, the temperature range over which the native state is stable ($T_G - T'_G$) is around 100 K. However, there are two exceptions, α -amylase ($T_G - T'_G \sim 53$ K) and MBP ($T_G - T'_G \sim 57$ K). Consistent with their large sizes, both have high values for ΔH_G (large slope) and ΔC_p (large curvature). Consequently, their stability curves are narrower (Figure 2).

The stabilities of proteins at the living temperatures of their source organisms are usually smaller than the maximal protein stabilities. Among homologous proteins, protein stabilities at the respective source organism living temperatures remain relatively constant (8). The proteins we studied derive from mesophilic and (hyper)thermophilic sources

Table 3: Derived Thermodynamic Characteristics for 26 Unique Proteins in Our Database

protein	ΔS_G (cal mol ⁻¹ K ⁻¹)	T_G^* (K)	T_S (K)	$\Delta G(T_S)$ (kcal/mol)	$\Delta G(298\text{ K})$ (kcal/mol)
α -amylase	751.1	264	290	10.3	9.3
Sso7d	170.9	201	282	8.0	7.7
TmCsp	176.3	253	303	4.8	4.7
BsHPr	167.7	256	300	4.0	4.0
λ -repressor ₆₋₈₅	205.8	244	286	4.6	4.3
barstar	209.3	244	293	5.7	5.6
Rnase Sa	302.9	209	263	9.1	5.8
Rnase T ₁	296.5	216	267	8.4	5.6
Rnase A	305.4	186	257	12.5	8.8
binase	358.3	120	217	21.8	10.3
MBP	718.1	279	307	10.7	9.6
OBP	273.1	180	257	12.2	9.2
ROP	402.7	243	292	10.8	10.6
Snase	294.5	246	285	6.2	5.5
FeCyt <i>b</i> ₅₆₂	276.2	268	303	5.2	5.1
Arc repressor	217.0	246	286	4.6	4.2
GDH domain II	204.8	252	296	4.9	4.9
H2A-H2B dimer	184.3	236	285	4.9	4.6
C _L fragment	199.0	266	299	3.5	3.5
TTF-1HD	80.4	188	249	2.8	1.3
STI	308.6	259	295	5.9	5.8
apoflavodoxin	191.0	245	287	4.3	4.0
thioredoxin	296.8	246	301	9.0	9.0
MetJ dimer	369.9	225	274	9.9	7.8
ASAI	512.7	247	292	12.5	12.3
ADA2h	135.9	251	299	3.6	3.6
mean	292.7 \pm 159.5	233 \pm 35	283 \pm 21	7.7 \pm 4.2	6.4 \pm 2.7
range	80.4-751.1	120.4-278.6	217-307	2.8-21.8	1.3-12.3
median	274.7	245	288	6.0	5.6

where the living temperatures of the source organisms (T_L) are invariably greater than T_S , the temperature of maximal protein stability. Similarly, α -amylase from the psychrophilic bacterium *A. haloplanctis* has a lower stability [$\Delta G(T_L) = 6.0$ kcal/mol] at the T_L (273K) of the bacterium than the maximal protein stability [$\Delta G(T_S) = 10.3$ kcal/mol] at T_S (290 K). Table 3 presents the thermodynamic characteristics of the 26 unique proteins. The average temperature of maximal protein stability is 283 ± 21 K. This observation is similar to that of Rees and Robertson (19). Figure 3 plots the distribution of the temperatures of maximal stability (T_S) for the 26 unique proteins. The distribution can be divided into two distinct categories. The first category contains 20 ($\sim 77\%$) unique proteins. All are maximally stable at temperatures above 273 K (range of 274–307 K, $T_S^{\text{av}} = 293 \pm 8$ K). Consistently, their stabilities at room temperature [$\Delta G(298\text{ K})$] are similar to their maximal stabilities [$\Delta G(T_S)$]. Proteins in the second category are maximally stable at temperatures away from room temperature. There are six ($\sim 23\%$) unique proteins in this category, all with T_S values below 273 K (range of 217–267 K, $T_S^{\text{av}} = 252 \pm 18$ K) (Table 3 and Figure 3). Here, protein stabilities at room temperature differ from the maximal protein stabilities (Table 3). Below, we describe the characteristics of the proteins in these two categories.

Proteins with Maximal Stabilities around Room Temperature. Among the 20 unique proteins with average temperatures of maximal stability (T_S^{av}) around room temperature, α -amylase is from a psychrophile and SSo7d, TmCsp, and TmGDH domain II are from thermophiles. The remaining 16 proteins are from mesophiles. Recently, we have compared the thermodynamic properties among five families of homologous thermophilic and mesophilic proteins (8). In that analysis, too, both thermophilic and mesophilic proteins were

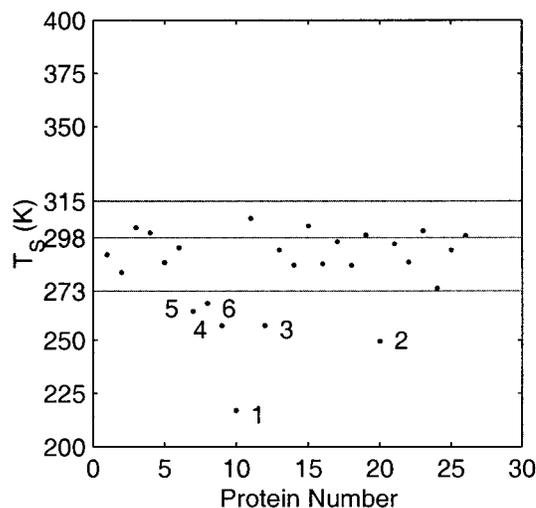


FIGURE 3: Distribution of the temperature of maximal stability (T_S) for 26 unique proteins in our database. The X-axis denotes the protein number, and the Y-axis denotes T_S . Three horizontal lines denote 273 (freezing point of water), 298 (room temperature), and 315 K. Most proteins in our database have T_S values around room temperature. Twenty of the 26 unique proteins have T_S values within the range of 273–315 K. The protein stability curves of these 20 unique proteins are shown in panels a and b of Figure 1. Six unique proteins have T_S values of < 273 K. These proteins are numbered 1–6 in the plot. These proteins are (1) binase, (2) TTF-1HD, (3) OBP, (4) Rnase A, (5) Rnase Sa, and (6) Rnase T₁. The protein stability curves of these six proteins are shown in Figure 2c.

maximally stable around room temperature. Hence, it appears that the majority of reversible two-state folding proteins are maximally stable around room temperature irrespective of the optimum living temperatures of their source organisms.

The 20 unique proteins have different sizes (number of residues), amino acid sequences, and three-dimensional

structures. Of these 20 unique proteins, 13 are monomers with single domains (Table 1 and Figure 2a). The remaining seven proteins contain two domains. Two of these seven are monomers, and five are dimers with a single domain per subunit (Figure 2b). The average temperature of maximal stability (T_S^{av}) for these 13 single-domain proteins is 295 ± 7 K (22 ± 7 °C). The T_S^{av} for the remaining seven is 289 ± 10 K (16 ± 10 °C). Hence, the maximal stability temperatures of the single-domain proteins are slightly closer to room temperature than those of the two-domain proteins. This observation is also supported by visual examination of panels a and b of Figure 2. The differences in the average temperatures of maximal stability between single- and two-domain proteins may be interpreted in terms of the differences in protein architecture. A recent hydrophobic effect-based computational procedure for protein folding describes a protein structure as consisting of a combination of hydrophobic folding units (HFUs). An HFU has a sufficiently large buried hydrophobic core and is capable of an independent thermodynamically stable existence (27, 28, 30). HFUs may (or may not) coincide with the structural domains in a protein. Twelve of the 13 single-domain proteins in our database for which structures are available contain a single cooperative HFU (Table 2). The 13th [*T. maritima* cold shock protein (TmCsp)] also most likely contains a single HFU as indicated by the structures of the homologous cold shock proteins from *E. coli*, *B. subtilis*, and *Bacillus caldolyticus*. On the other hand, five of the seven two-domain proteins contain two HFUs. The psychrophilic α -amylase contains three HFUs, and the Arc repressor dimer contains a single HFU spanning the interface (Table 2). The presence of more than one cooperative HFU in most of the two-domain proteins may have affected the stability curves of these proteins.

Proteins with Maximal Stabilities Away from Room Temperature. Six unique proteins have their maximal stabilities below 273 K (Rnase Sa, Rnase T₁, Rnase A, TTF-1HD, OBP, and binase). Their T_S^{av} is 252 ± 18 K (-21 ± 18 °C). Binase has significantly different values for all thermodynamic parameters, including T_S , as compared to the other proteins in our database (Table 3).

The accuracy in the estimate of the T_S value of a protein depends on the accuracies in the estimates of T_G , ΔH_G , and ΔC_p (eq 2). T_G values are usually accurate to better than $\pm 1\%$, and the value for ΔH_G can be determined to approximately $\pm 5\%$ (31). However, there may be considerable differences in ΔC_p values (31). While in general ΔC_p is taken to be independent of T , Privalov (2) has shown that it decreases slightly at low and high temperatures. In our database, the errors in estimates of T_G and ΔH_G are quite small, wherever such data are available. This is not always the case for ΔC_p . The error estimates for ΔC_p values are available for 24 of the 31 proteins in our database (Table 1). Nineteen of these 24 are unique proteins. For 16 of the 24 proteins, the errors in the ΔC_p values are within 10%. For five proteins (barstar, MBP, ROP, ferricytochrome *b*₅₆₂, and TmGDH domain II), the estimated errors in ΔC_p values are in the range of 10–21%. A visual examination of the stability curves (Figure 2) of these 21 proteins shows that the errors in the ΔC_p values do not significantly affect the values of T_S . These 21 proteins include 14 of the 20 unique

proteins with T_S values around room temperature and three (Rnase Sa, Sa2, and Sa3) of the six unique proteins with T_S values away from room temperature. The errors in the estimates of ΔC_p are greater than 25% for the remaining three of the 24 proteins [27.1% for binase, 34.7% for OBP, and 38.4% for ADA2h (Table 1)]. All three are unique proteins. The large uncertainties in their ΔC_p values significantly affect their T_S values (Figure 2). Two of these proteins (binase and OBP) are among the six unique proteins with T_S values below 273 K (Figure 2c). The third protein, ADA2h, is maximally stable at 299 K.

In the literature, the reported ΔC_p values for Rnase A lie in the range of 1.0–2.3 kcal mol⁻¹ K⁻¹ (23, 32). In our work, we have taken the ΔC_p of Rnase A to be 1.15 ± 0.09 , as suggested by Pace et al. (32). However, if we had taken a literature average value of 2.0 ± 0.24 for the ΔC_p of Rnase A (Table 1 in ref 32), then the T_S of Rnase A would have been 287.5 K, with the same values of T_G and ΔH_G .

Thyroid transcription factor 1 homeodomain (TTF-1HD) appears to have lower hydrophobicity (66%, Table 2). The hydrophobicity essentially measures the extent to which the nonpolar surface of a protein is buried in its core. On average, its value is in the range of 75–85%. Hence, the hydrophobic effect appears to be weaker for TTF-1HD. In the original paper reporting the thermodynamic data for this protein, Damante et al. (33) have noted the structural disorder in TTF-1HD, particularly in the recognition helix C-terminal region. The protein stability curve of TTF-1HD (Figure 2c) and the thermodynamic stability data presented in Tables 1 and 2 indicate low stability for this protein at room temperature. This was also reported by the original investigators (33). Damante et al. (33) have further reported a value of 0.08 kcal mol⁻¹ K⁻¹ for both ΔC_p and ΔS_G . We found that this value is correct for ΔS_G by dividing ΔH_G (25.41 kcal/mol) by T_G (316 K). Hence, we have calculated the value of ΔC_p for TTF-1HD by substituting values for ΔG , T_G , and ΔH_G from the original report (33) into the Gibbs–Helmholtz equation. For this purpose, we have used the ΔG value (1.26 kcal/mol) determined by Damante et al. (33) from urea denaturation experiments at 25 °C. The calculated value of ΔC_p for TTF-1HD is 0.34 kcal mol⁻¹ K⁻¹. We have used this value of ΔC_p to plot the stability curve for TTF-1HD.

We do not have an adequate explanation for the remaining two proteins, Rnase Sa and Rnase T₁. At the same time, we also notice that the T_S values for these proteins [267 K (-6 °C) for Rnase T₁ and 263 K (-10 °C) for Rnase Sa] are closer to the 273 K mark (Table 3 and Figure 3). Rnase Sa has a hydrophobicity of 75% and Rnase T₁ 82%. Both proteins contain single hydrophobic folding units (HFUs) (Table 2). The estimated error in ΔC_p for Rnase Sa is $\sim 6\%$ (Table 1, ΔC_p for Rnase Sa = 1.52 ± 0.09 kcal mol⁻¹ K⁻¹) and does not affect its T_S value (Figure 2c). The data on the error in ΔC_p for Rnase T₁ (1.59) are not available in Table 1. Pace et al. (31) have estimated the energetic contributions of the hydrophobic effect and the intramolecular hydrogen bonding toward the conformational stabilities of Rnase Sa, Rnase Sa2, Rnase Sa3, Rnase T₁, barnase, and Rnase A (Table 5 in their paper). Both factors contribute significantly (with roughly equal magnitudes) to the stabilities of these ribonucleases. This evidence appears to indicate a relatively lesser role for the hydrophobic effect in the folding of these proteins.

Table 4: Average Characteristics of 12 Single-Domain Proteins

parameter	mean value	range	median value
N_{res}	113 ± 40	62–181	105
T_G (K)	343 ± 14	326–371	342
T'_G (K)	248 ± 17	201–268	249
T_S (K)	294 ± 8	282–303	296
ΔH_G (kcal/mol)	77.0 ± 17.6	58.1–106.9	69.1
ΔC_p (kcal mol ⁻¹ K ⁻¹)	1.58 ± 0.58	0.62–2.60	1.42
ΔS_G (cal/mol)	225.1 ± 53.1	167.7–308.6	205.3
$\Delta G(T_S)$ (kcal/mol)	5.5 ± 1.6	3.5–9.0	5.1
$\Delta G(298\text{ K})$ (kcal/mol)	5.3 ± 1.6	3.5–9.0	5.0

Table 5: Correlations among Various Parameter Pairs in 12 Single-Domain Proteins^a

parameter pair	r	t value	parameter pair	r	t value
T_G and Δh_G	0.69	3.01	N_{res} and Δh_G	-0.80	-4.22
T_G and Δs_G	0.59	2.31	N_{res} and Δs_G	-0.79	-4.07
T_G and ΔC_p	-0.63	-2.57	N_{res} and $\Delta g(T_S)$	-0.71	-3.19
T_G and $\Delta g(T_S)$	0.86	5.33	N_{res} and ΔC_p	0.62	2.50
Δh_G and $\Delta g(T_S)$	0.84	4.90	ΔH_G and ΔC_p	0.78	3.94
Δs_G and $\Delta g(T_S)$	0.78	3.94	ΔS_G and ΔC_p	0.85	5.10

^a The significance of a correlation for a parameter pair is given by its t value. The procedure for computing the t values is described in Materials and Methods. The correlations for all parameter pairs shown here are significant at the $\geq 95\%$ level of confidence. The correlations for parameter pairs shown in bold are significant at the 99.5% level of confidence. Values of Δh_G , Δs_G , and $\Delta g(T_S)$ were obtained by normalizing ΔH_G , ΔS_G , and $\Delta G(T_S)$, respectively, by the number of residues (N_{res}).

Taken together, these observations indicate that proteins may have maximal stability temperatures (T_S) away from room temperature due to uncertainties in their ΔC_p values, a relatively weaker hydrophobic effect, and/or stronger electrostatic interactions.

Correlations among Various Thermodynamic Parameters. The finding that most reversible two-state proteins are maximally stable around room temperature provides an opportunity to examine whether the variations in the thermodynamic parameters across different proteins are correlated. Analysis of such correlations bears on aspects of protein folding and stability and on thermal adaptation of the proteins to the living temperature of their respective source organism. We have performed linear regression among various thermodynamic parameters for the single-domain proteins that are maximally stable around room temperature. The thermodynamic formalism of the reversible two-state $N \rightleftharpoons D$ transition applies most adequately to proteins that contain a single hydrophobic cooperative folding \rightleftharpoons unfolding unit.

We have taken 12 of the 13 single-domain proteins whose protein stability curves are shown in Figure 2a. ADA2h has not been included in this analysis because of a large error in its ΔC_p value (Table 1 and previous section). Table 4 presents the average characteristics of these 12 single-domain proteins, and Table 5 summarizes the biologically meaningful and statistically significant correlations observed among the thermodynamic parameters of these proteins. The significance of the correlation among any two parameters is measured by the t test (see Materials and Methods). All the parameter pairs listed in Table 5 exhibit correlations that are significant at least at the 95% level of confidence ($r \geq 0.5$, $p < 5 \times 10^{-2}$). For the parameter pairs shown in bold, the

correlation is significant at the 99.5% level of confidence ($r \geq 0.7$, $p < 5 \times 10^{-3}$).

Consistent with our observations described above, there are no significant correlations of T_S with N_{res} (number of residues) or with the various thermodynamic parameters. Table 5 illustrates that T_G is positively correlated with Δh_G , Δs_G , and $\Delta g(T_S)$. Δh_G , Δs_G , and $\Delta g(T_S)$ were obtained by normalizing ΔH_G , ΔS_G , and $\Delta G(T_S)$, respectively, by N_{res} for each protein. The normalization removes artifacts due to the differences in the protein size. Δh_G and Δs_G are positively correlated with $\Delta g(T_S)$. We also observe a negative correlation between T_G and ΔC_p . The correlation coefficients for the parameter pairs T_G and Δh_G ($r = 0.69$) and T_G and ΔC_p ($r = -0.63$) are significant at the 95% level of confidence but not at the 99.5% level. Panels a and b of Figure 4 plot these parameter pairs. The correlations between the parameter pairs T_G and $\Delta g(T_S)$ ($r = 0.86$) and Δh_G and $\Delta g(T_S)$ ($r = 0.84$) (panels c and d of Figure 4, respectively) are significant at the 99.5% level of confidence. The plots in panels c and d of Figure 4 for the parameter pairs T_G and $\Delta g(T_S)$ and Δh_G and $\Delta g(T_S)$ show one outlier each, the hyperthermophilic protein Sso7d. After Sso7d had been removed, the linear correlation coefficients for the two parameter pairs are 0.76 and 0.92, respectively. Taken together, these correlations may have significant implications for the temperature adaptation shown by thermophilic proteins (see the Discussion).

The protein size (N_{res}) shows strong negative correlations with Δh_G , Δs_G , and $\Delta g(T_S)$ (Table 5 and Figure 5). All these negative correlations are significant at the 99.5% level of confidence. The negative correlation between N_{res} and $\Delta g(T_S)$ is also significant for the seven two-domain proteins whose stability curves are shown in Figure 2b. These results indicate that there may be an upper limit to maximal protein stability independent of the protein size. Consistently, we observe no significant correlation between N_{res} and $\Delta G(T_S)$.

DISCUSSION

Here, our premise is simple. It is based on two main pieces of evidence. First, the hydrophobic effect is the major force driving protein folding (14, 15). Second, the solubilities of small organic solutes and hydrophobic amino acids in water are minimal at (around) room temperature, and the hydrophobic effect is the strongest in this temperature range (16, 17). Hence, we can expect that globular proteins with reasonably large hydrophobic cores should be maximally stable around room temperature. Nonhydrophobic, e.g., electrostatic and disulfide interactions, may also contribute significantly toward protein stability (34). The relative extents to which the hydrophobic effect and the electrostatic interactions contribute to protein stability are controversial. According to Pace et al. (12, 13, 31), both hydrophobic interactions and intramolecular hydrogen bonding make large but comparable contributions to protein stability. On the other hand, a simple calculation by Schellman (16) has shown that approximately 75% of protein stability at room temperature derives from the hydrophobic effect. Clearly, the relative strengths of these interactions are likely to vary with the protein sequence. Hence, we cannot expect to see an exact coincidence between room temperature and the T_S values for proteins. We find that 20 of the 26 unique proteins in our database are maximally stable around room temperature.

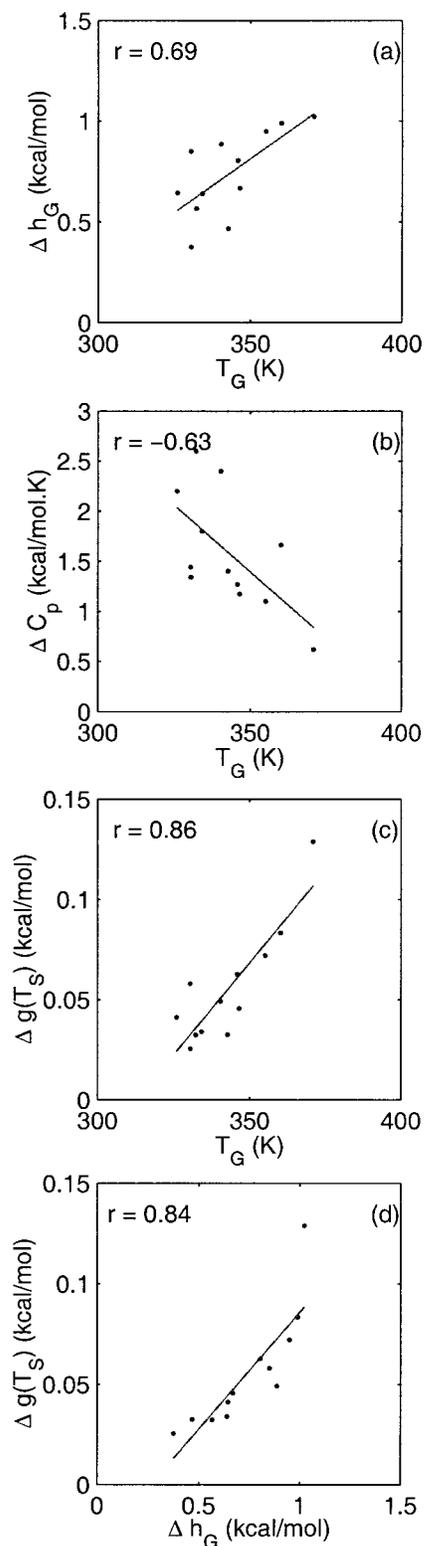


FIGURE 4: Plots showing correlations among the parameter pairs (a) T_G and Δh_G , (b) T_G and ΔC_p , (c) T_G and $\Delta g(T_S)$, and (d) Δh_G and $\Delta g(T_S)$ in 12 single-domain proteins which are maximally stable around room temperature. For each protein, values of Δh_G and $\Delta g(T_S)$ were computed by normalizing ΔH_G and $\Delta G(T_S)$ by the number of residues (N_{res}) in the protein, respectively. In each plot, a least-squares line obtained by regression analysis is shown and the linear correlation coefficient (r) is indicated in the upper left corner. In plots c and d, there is an outlier due to high values of $\Delta g(T_S)$, T_G , and Δh_G for the hyperthermophilic protein Sso7d. After Sso7d had been removed from the regression analysis, the linear correlation coefficients for the rest of the 11 proteins are 0.76 for T_G and $\Delta g(T_S)$ and 0.92 for Δh_G and $\Delta g(T_S)$.

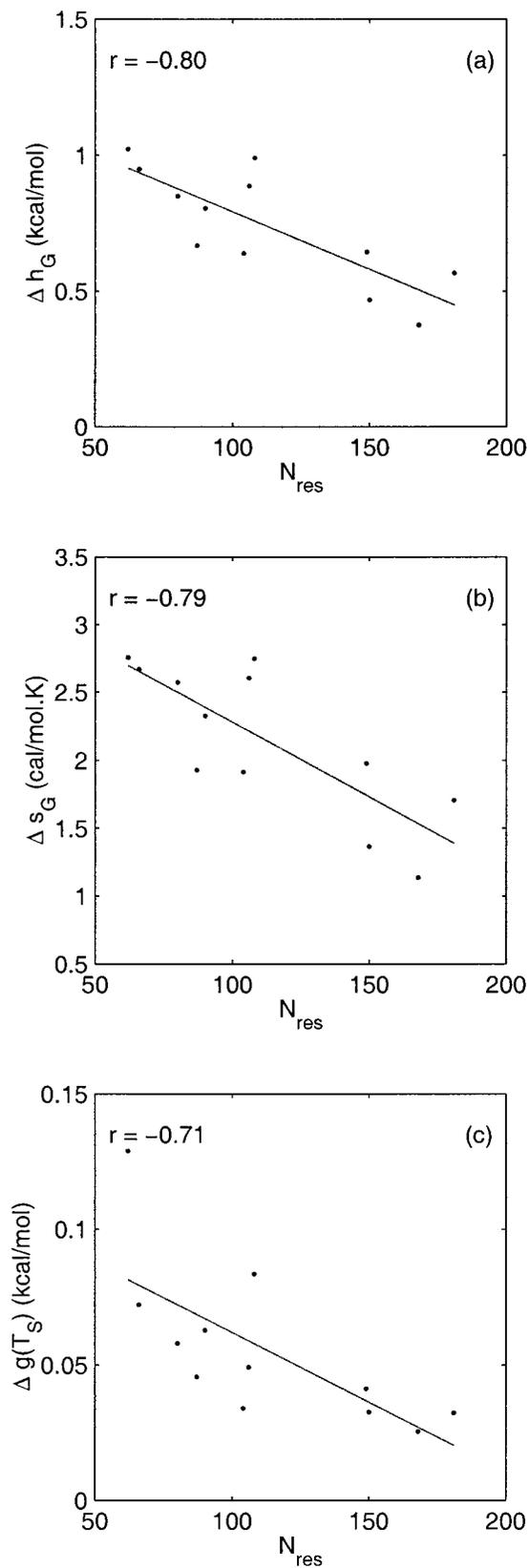


FIGURE 5: Plots showing correlations among the parameter pairs (a) N_{res} and Δh_G , (b) N_{res} and ΔS_G , and (c) N_{res} and $\Delta g(T_S)$ in 12 single-domain proteins which are maximally stable around room temperature. For each protein, the values of Δh_G , ΔS_G , and $\Delta g(T_S)$ were computed by normalizing ΔH_G , ΔS_G , and $\Delta G(T_S)$ by the number of residues (N_{res}) in the protein, respectively. In each plot, a least-squares line obtained by regression analysis is shown and the linear correlation coefficient (r) is indicated in the upper left corner.

For six unique proteins, the T_S values are away from room temperature. Two of these proteins have large uncertainties in their estimates of ΔC_p values. Recently, Loladze et al. (35) have also emphasized that ΔC_p is the key parameter in protein thermodynamic studies. According to Pace and his colleagues, the best estimate of the ΔC_p value of a protein can only be obtained by a global analysis of both chemical and thermal denaturation data (31). However, few original experimental reports indicate that such a global analysis has been performed. Apart from these two proteins, another protein (TTF-1HD) has a weaker hydrophobic effect and high local structural disorder along with an unreliable reported value of ΔC_p . The value of ΔC_p is also controversial for Rnase A (32). For the two remaining proteins, the T_S values are considerably higher. In these proteins, the hydrophobic effect and intramolecular hydrogen bonding interactions make large but almost equal contributions (31). Here we cannot rule out the possibility that proteins with T_S values away from room temperature may have a weaker hydrophobic effect and/or stronger electrostatic contributions.

Each protein can be considered an independent thermodynamic system. Therefore, there may not be any correlation among the thermodynamic parameters of two different proteins, unless they share a common property. Maximal stability around room temperature appears to be such a property. Such an argument is substantiated by evidence. The temperature of maximal protein stability is independent of heat and of cold denaturation temperatures, of the living temperatures of the source organism, and of the sequence and structural properties of the proteins in our database. Due to the common maximal stability temperature, the correlations observed among the various thermodynamic properties of the 12 single-domain proteins (Figures 4 and 5) may be biologically meaningful.

In principle, to derive the correlations among various thermodynamic parameters for the proteins, it is essential that they be determined under identical experimental conditions. However, this is infeasible. Experiments are performed in different laboratories under different conditions. In our effort described here, we have been able to collect a highly homogeneous database (see Data Collection in Materials and Methods) with a barely sufficient number of proteins to perform a statistically meaningful analysis. Furthermore, the differences in the denatured states of the proteins may affect the observed values of the correlation coefficients among the thermodynamic parameters. The parameters $\Delta G(T_S)$, ΔH_G , and ΔC_p are described in terms of the differences between the native and denatured states of the proteins, yet the denatured states are not extended random coils. They may have a considerable extent of residual structural content with some nativelike topology (36). This may affect the experimental values, especially ΔC_p since it is related to the nonpolar surface area buried upon protein folding (37–39). In a database containing unique proteins, these errors are likely to be random since the denatured states of different proteins will have different residual structural contents. Unfortunately, there is no way of accounting for the effect that these errors may have. It is, therefore, remarkable that we observe the correlations despite these potential sources of errors.

In our previous statistical analysis of 19 homologous thermophilic and mesophilic two-state proteins in five

different families (8), we did not observe a correlation between T_G and ΔC_p . However, the thermophilic proteins had lower ΔC_p values than their mesophilic homologues in three of the five families. In one family, the ΔC_p value for the thermophilic protein was similar to that of its mesophilic homologues. In the fifth family, the ΔC_p values for thermophilic and mesophilic proteins were not experimentally determined (8). Statistical differences in the data used in the previous and present studies, the availability of limited data for both studies, and the relatively larger uncertainties in experimental ΔC_p values may be responsible for this difference between our two studies. Furthermore, we have been able to uphold higher-quality thermodynamic data in the study presented here (see Materials and Methods).

The correlations seen in the 12 unique single-domain proteins yield a consistent picture with regard to protein temperature adaptation. Previously, three potential mechanisms for achieving a greater temperature resistance in proteins have been suggested (19, 40–43). First, maximal protein stability [$\Delta G(T_S)$] is greater for thermophilic proteins, resulting in an upshift of their protein stability curves. Second, thermophilic proteins may have $\Delta G(T_S)$ values similar to those of their mesophilic homologues, but their protein stability curves may be broader. Third, we may observe a simple left to right shift between the protein stability curves of the thermophilic and mesophilic proteins. Any mechanism must satisfy the condition that protein stability at the respective organism living temperature remain almost constant (8). Our current results suggest that the third mechanism may not be favored if it involves a significant deviation of the protein T_S away from room temperature. The positive correlations among T_G , Δh_G , and $\Delta g(T_S)$ and the negative correlation between T_G and ΔC_p suggest that a combination of the first and second mechanisms may be favored. Hence, proteins may achieve greater temperature resistance by broadening and upshifting their protein stability curves such that their temperatures of maximal stability still remain around room temperature. Consistently, in our previous analysis, we observed that protein stability curves of thermophilic proteins are upshifted and broader than those of their mesophilic homologues (8).

These observations are consistent with studies on thermophilic proteins (44, 45). Analyses of derived amino acid sequences of the completed genomes of (hyper)thermophilic organisms indicate that these proteins tend to be shorter (refs 44 and 45 and references therein). A smaller number of residues indicates lower ΔC_p values for the thermophilic proteins. At the same time, both the proteome-wide comparisons of thermophilic and mesophilic proteins and the sequence and structural comparisons of families containing homologous thermophilic and mesophilic proteins show a greater occurrence of charged residues in the thermophilic proteins. This yields a larger number of electrostatic interactions such as salt bridges and their networks (8, 44–46). Formation of these specific interactions would result in increased ΔH_G values for the thermophilic proteins (8, 45). Hence, while the hydrophobic effect is dominant in the folding of both thermophilic and mesophilic proteins, the contribution due to electrostatic interactions is significant for the thermophile–mesophile protein stability differentials.

Our results also show that the enthalpic and entropic change at the melting temperature as well as the maximal

protein stability per amino acid decrease with an increase in protein size (Figure 5). This suggests an upper limit for maximal protein stability. In general, maximal protein stabilities $\Delta G(T_S)$ tend to be on the order of 10 kcal/mol (31, 47). We observe the same range for proteins in our database, irrespective of size. A positive correlation between N_{res} and ΔC_p is well-known for proteins (37, 39). A greater value of ΔC_p would result in a smaller value for $\Delta G(T_S)$, since $\Delta G(T_S)$ is given by (1, 2)

$$\Delta G(T_S) = \Delta H_G - (T_G - T_S)\Delta C_p \quad (7)$$

Consistently, Liang and Dill (48) have illustrated that larger proteins tend to be more loosely packed than smaller ones. Considerations of function may require such an upper limit to $\Delta G(T_S)$ values for proteins.

CONCLUSIONS

Consistent with the hydrophobic effect being both the major force in protein folding and the strongest at room temperature, we observe that more than three-fourths of the proteins in our database are maximally stable around room temperature. In the remaining proteins, the temperatures of maximal protein stability deviate from room temperature due to uncertainties in the estimated values of ΔC_p , a weaker hydrophobic effect, and/or stronger electrostatic interactions. The requirement of keeping the maximal stability around room temperature results in correlated variations in the thermodynamic parameters across different proteins. These correlations suggest a possible mechanism for thermal adaptation of proteins to the living temperatures of their source organism. Further, there may be an upper limit for maximal protein stability, regardless of protein size.

ACKNOWLEDGMENT

We thank Drs. Buyong Ma, Neeti Sinha, Gunasekaran Kannan, David Zanuy, and, in particular, Jacob V. Maizel for numerous helpful discussions. The content of this publication does not necessarily reflect the view or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the U.S. Government.

REFERENCES

- Becktel, W., and Schellman, J. A. (1987) Protein stability curves, *Biopolymers* 26, 1859–1877.
- Privalov, P. L. (1990) Cold denaturation of proteins, *Crit. Rev. Biochem. Mol. Biol.* 25, 281–305.
- Griko, Y. V. (2000) Energetic basis of structural stability in the molten globule state: α -lactalbumin, *J. Mol. Biol.* 297, 1259–1268.
- Hallerbach, B., and Hinz, H. J. (1999) Protein heat capacity: inconsistencies in the current view of cold denaturation, *Biophys. Chem.* 76, 219–227.
- Brandts, J. F. (1964) *J. Am. Chem. Soc.* 86, 4291–4314.
- Vedamuthu, M., Singh, S., and Robinson, G. W. (1994) Properties of liquid water: origin of the density anomalies, *J. Phys. Chem.* 98, 2222–2230.
- Robinson, G. W., and Cho, C. H. (1999) Role of hydration water in protein unfolding, *Biophys. J.* 77, 3311–3318.
- Kumar, S., Tsai, C. J., and Nussinov, R. (2001) Thermodynamic differences among homologous thermophilic and mesophilic proteins, *Biochemistry* 40, 14152–14165.
- Tsai, C. J., Maizel, J. V., and Nussinov, R. (2002) The hydrophobic effect: A new insight from cold denaturation and a two-state water structure, *Crit. Rev. Biochem. Mol. Biol.* (in press).
- Ruelle, P., and Kesselring, U. W. (1998) The hydrophobic effect. I. A consequence of the mobile order in H-bonded liquids, *J. Pharm. Sci.* 87, 987–997.
- Klotz, I. M. (1999) Parallel change with temperature of water structure and protein behavior, *J. Phys. Chem. B* 103, 5910–5916.
- Pace, C. N., Shirley, B. A., McNutt, M., and Gajiwala, K. (1996) Forces contributing to the conformational stability of proteins, *FASEB J.* 10, 75–83.
- Pace, C. N. (2001) Polar group burial contributes more to protein stability than nonpolar group burial, *Biochemistry* 40, 310–313.
- Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation, *Adv. Protein Chem.* 14, 1–63.
- Dill, K. A. (1990) Dominant forces in protein folding, *Biochemistry* 29, 7133–7155.
- Schellman, J. A. (1997) Temperature, stability and hydrophobic interaction, *Biophys. J.* 73, 2960–2964.
- Privalov, P. L., and Gill, S. J. (1988) Stability of protein structure and hydrophobic interaction, *Adv. Protein Chem.* 39, 191–234.
- Fasman, G. D., Ed. (1976) *Handbook of biochemistry and molecular biology*, 3rd ed., pp 114–115, CRC Press, Cleveland.
- Rees, D. C., and Robertson, A. D. (2001) Some thermodynamic implications for the thermostability of proteins, *Protein Sci.* 10, 1187–1194.
- Robertson, A. D., and Murphy, K. P. (1997) Protein structure and energetics of protein stability, *Chem. Rev.* 97, 1251–1267.
- Ganesh, C., Eswar, N., Srivastava, S., Ramakrishnan, C., and Varadarajan, R. (1999) Prediction of the maximal stability temperature of monomeric globular proteins solely from amino acid sequence, *FEBS Lett.* 454, 31–36.
- Gromiha, M. M., An, J., Kono, H., Oobatake, M., Uedaira, H., Prabhakaran, P., and Sarai, A. (2000) ProTherm, version 2.0: thermodynamic database for proteins and mutants, *Nucleic Acids Res.* 28, 283–285.
- Pfeil, W. (1998) *Protein stability and folding. A collection of thermodynamic data*, Springer-Verlag, Heidelberg, Germany.
- Privalov, P. L., and Khechinashvili, N. (1974) A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study, *J. Mol. Biol.* 86, 665–684.
- Privalov, P. L. (1979) Stability of proteins, *Adv. Protein Chem.* 33, 167–241.
- Bernstein, F. C., Koetzle, T. F., Williams, G. J., Myer, E. E., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) The Protein Data Bank: A computer-based archival file for macromolecular structures, *J. Mol. Biol.* 112, 535–542.
- Tsai, C. J., and Nussinov, R. (1997) Hydrophobic folding units derived from dissimilar monomer structures and their interactions, *Protein Sci.* 6, 24–42.
- Tsai, C. J., and Nussinov, R. (1997) Hydrophobic folding units at protein–protein interfaces: Implications to protein folding and protein–protein association, *Protein Sci.* 6, 1426–1437.
- Spiegel, M. R. (1980) *Theory and problems of probability and statistics, SI (metric) edition, Schaum's outline series*, McGraw-Hill, New Delhi.
- Tsai, C. J., Ma, B., Kumar, S., Wolfson, H., and Nussinov, R. (2001) Protein Folding: Binding of conformationally fluctuating building blocks via population selection, *Crit. Rev. Biochem. Mol. Biol.* 36, 399–433.
- Pace, C. N., Hebert, E. J., Shaw, K. L., Schell, D., Both, V., Krajcikova, D., Sevcik, J., Wilson, K. S., Dauter, Z., Hartley, R. W., and Grimsley, G. R. (1998) Conformational stability and thermodynamics of folding of ribonucleases Sa, Sa2 and Sa3, *J. Mol. Biol.* 279, 271–286.

32. Pace, C. N., Grimsley, G. R., Thomas, S. T., and Makhatadze, G. I. (1999) Heat capacity change for ribonuclease A folding, *Protein Sci.* 8, 1500–1504.
33. Damante, G., Tell, G., Leonardi, A., Fogolari, F., Bartolotti, N., Di Lauro, R., and Formisano, S. (1994) Analysis of the conformation and stability of rat TTF-1 homeodomain by circular dichroism, *FEBS Lett.* 354, 293–296.
34. Kumar, S., and Nussinov, R. (1999) Salt bridge stability in monomeric proteins, *J. Mol. Biol.* 293, 1241–1255.
35. Loladze, V. V., Ermolenko, D. N., and Makhatadze, G. I. (2001) Heat capacity changes upon burial of polar and nonpolar groups in proteins, *Protein Sci.* 10, 1343–13452.
36. Shortle, D., and Ackerman, M. S. (2001) Persistence of native-like topology in a denatured protein in 8 M urea, *Science* 293, 487–489.
37. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding, *Protein Sci.* 4, 2138–2148.
38. Privalov, P. L., Tiktopulo, E. I., Venyaminov, S. Y., Griko, Y. V., Makhatadze, G. I., and Khechinashvili, N. N. (1989) Heat capacity and conformation of proteins in a denatured state, *J. Mol. Biol.* 205, 737–750.
39. Edgcomb, S. P., and Murphy, K. P. (2000) Structural energetics of protein folding and binding, *Curr. Opin. Biotechnol.* 11, 62–66.
40. McCrary, B. S., Edmondson, S. P., and Shriver, J. W. (1996) Hyperthermophile protein folding thermodynamics: Differential scanning calorimetry and chemical denaturation of Sac7d, *J. Mol. Biol.* 264, 784–805.
41. Beadle, B. M., Baase, W. A., Wilson, D. B., Gilkes, N. R., and Shoichet, B. K. (1999) Comparing the thermodynamic stabilities of a related thermophilic and mesophilic enzyme, *Biochemistry* 38, 2570–2576.
42. Alexander, P., Fahnestock, S., Lee, T., Orban, J., and Bryan, P. (1992) Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2. Why small proteins tend to have high denaturation temperatures, *Biochemistry* 31, 3597–3603.
43. Rees, D. C., and Adams, M. W. W. (1995) Hyperthermophiles: taking the heat and loving it, *Structure* 3, 251–254.
44. Sterner, R., and Liebl, W. (2001) Thermophilic adaptation of proteins, *Crit. Rev. Biochem. Mol. Biol.* 36, 39–106.
45. Kumar, S., and Nussinov, R. (2001) How do thermophilic proteins deal with heat? *Cell. Mol. Life Sci.* 58, 1216–1233.
46. Kumar, S., Tsai, C. J., and Nussinov, R. (2000) Factors enhancing protein thermostability, *Protein Eng.* 13, 179–191.
47. Pace, C. N. (1990) Conformational stability of globular proteins, *Trends Biochem. Sci.* 15, 14–17.
48. Liang, J., and Dill, K. A. (2001) Are proteins well packed? *Biophys. J.* 81, 751–766.
49. Feller, G., d'Amico, D., and Gerday, C. (1999) Thermodynamic stability of a cold-active α -amylase from the antarctic bacterium *Alteromonas haloplacis*, *Biochemistry* 38, 4613–4619.
50. Knapp, S., Karshikoff, A., Berndt, K. D., Christova, P., Atanasov, B., and Ladenstein, R. (1996) Thermal unfolding of the DNA-binding protein Sso7d from the hyperthermophile *Sulfolobus solfataricus*, *J. Mol. Biol.* 264, 1132–1144.
51. Knapp, S., Mattson, P. T., Christova, P., Berndt, K. D., Karshikoff, A., Vihinen, M., Smith, C. I. E., and Ladenstein, R. (1998) Thermal unfolding of small proteins with SH3 domain folding pattern, *Proteins: Struct., Funct., Genet.* 31, 309–319.
52. Wassenberg, D., Welker, C., and Jaenicke, R. (1999) Thermodynamics of the unfolding of the cold shock protein from *Thermotoga maritima*, *J. Mol. Biol.* 289, 187–193.
53. Scholtz, J. M. (1995) Conformational stability of HPr: The histidine-containing phosphocarrier protein from *Bacillus subtilis*, *Protein Sci.* 4, 35–43.
54. Nicholson, E. M., and Scholtz, J. M. (1996) Conformational stability of *Escherichia coli* HPr protein: Test of the linear extrapolation method and a thermodynamic characterization of cold denaturation, *Biochemistry* 35, 11369–11378.
55. Huang, G. S., and Oas, T. G. (1996) Heat and cold denatured states of monomeric λ repressor are thermodynamically and chemically equivalent, *Biochemistry* 35, 6173–6180.
56. Schoppe, A., Hinz, H. J., Agashe, V. R., Ramachandran, S., and Udgaonkar, J. B. (1997) DSC studies of the conformational stability of barstar wild-type, *Protein Sci.* 6, 2196–2202.
57. Hu, C. Q., Strutevant, J. M., Thomson, J. A., Erickson, R. E., and Pace, C. N. (1992) Thermodynamics of ribonuclease T₁ denaturation, *Biochemistry* 31, 4876–4882.
58. Protasevich, I. I., Platonov, A. L., Pavlovsky, A. G., and Esipova, N. G. (1987) Distribution of charges in *Bacillus intermedius* 7P ribonuclease determines the number of cooperatively melting regions of the globule, *J. Biomol. Struct. Dyn.* 4, 885–893.
59. Ganesh, C., Shah, A. N., Swaminathan, C. P., Surolia, A., and Varadarajan, R. (1997) Thermodynamic characterization of the reversible, two state unfolding of maltose binding protein, a large two domain protein, *Biochemistry* 36, 5020–5028.
60. Burova, T. V., Choiset, Y., Jankowski, C. K., and Haertle, T. (1999) Conformational stability and binding properties of porcine odorant binding protein, *Biochemistry* 38, 15043–15051.
61. Steif, C., Hinz, H.-J., and Cesareni, G. (1995) Effects of cavity creating mutations on conformational stability and structure of the dimeric 4- α -helical protein ROP: Thermal unfolding studies, *Proteins: Struct., Funct., Genet.* 23, 83–96.
62. Shortle, D., Meeker, A. K., and Freire, E. (1988) Stability mutants of staphylococcal nuclease: Large compensating enthalpy–entropy changes for the reversible denaturation reaction, *Biochemistry* 27, 4761–4768.
63. Feng, Y., and Sligar, S. G. (1991) Effect of heme binding on the structure and stability of *Escherichia coli* apocytochrome *b*₅₆₂, *Biochemistry* 30, 10150–10155.
64. Bowie, J. U., and Sauer, R. T. (1989) Equilibrium dissociation and unfolding of the Arc repressor dimer, *Biochemistry* 28, 7139–7143.
65. Consalvi, V., Chiaraluce, R., Giangiacomo, L., Sandurra, R., Christova, P., Karshikoff, A., Knapp, S., and Ladenstein, R. (2000) Thermal unfolding and conformational stability of recombinant domain II of glutamate dehydrogenase from the hyperthermophile *Thermotoga maritima*, *Protein Eng.* 13, 501–507.
66. Karantz, V., Baxevanis, A. D., Freire, E., and Moudrianakis, E. N. (1995) Thermodynamic studies of the core histones: Ionic strength and pH dependence of H2A–H2B dimer stability, *Biochemistry* 34, 5988–5996.
67. Okajima, T., Kawata, Y., and Hamaguchi, K. (1990) Chemical modification of tryptophan residues and stability changes in proteins, *Biochemistry* 29, 9168–9175.
68. Fukada, H., Kitamura, S., and Takahashi, K. (1995) Calorimetric study of the thermal unfolding of kunitz-type soybean trypsin inhibitor at pH 7.0, *Thermochim. Acta* 266, 365–372.
69. Genzor, C. G., Beldarrain, A., Gomez-Moreno, C., Lopez-Lacomba, J. L., Cortijo, M., and Sancho, J. (1996) Conformational stability of apoflavodoxin, *Protein Sci.* 5, 1376–1388.
70. Santoro, M. M., and Bolen, D. W. (1992) A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range, *Biochemistry* 31, 4901–4907.
71. Johnson, C., Cooper, A., and Stockley, P. G. (1992) Differential scanning calorimetry of thermal unfolding of the methionine repressor protein (MetJ) from *Escherichia coli*, *Biochemistry* 31, 9717–9724.
72. Bachhawat, K., Kapoor, M., Dam, T. K., and Surolia, A. (2001) The reversible two-state unfolding of a monocot mannose-binding lectin from garlic bulbs reveals the dominant role of the dimeric interface in its stabilization, *Biochemistry* 40, 7291–7300.
73. Villegas, V., Azuaga, A., Catusas, L., Reverter, D., Mateo, P. L., Avelles, F. X., and Serrano, L. (1995) Evidence for a two state transition in the folding process of the activation domain of human procarboxypeptidase A2, *Biochemistry* 34, 15105–15110.